Part

B

Section 11: Logistic Regression Model of Detection Probability for Marbled Murrelets

Overview and Introduction

At the third meeting of the Science Advisory Panel, there was extensive discussion of the most appropriate metric for determining the relative numbers of Murrelets breeding in different forest stands. Several metrics were discussed (see text of HCP). One of these metrics was to use the ‘relative bird value’ of the stand, as proposed by Redwood Sciences lab.

After lengthy discussion, the Panel concluded that some of the assumptions of this method needed further testing, and depending on the results of these tests, modification. Until such time as these tests and modifications were complete, occupancy data on an acreage basis were the most relevant and best supported technique to assess Murrelet ‘take’. The Panel stated:

“Analysis should examine variation in occupancy as explained by variation in habitat attributes. It should explore variance in per-visit detection likelihood as a function of site attributes to estimate the probability of making a correct call on occupancy as a function of habitat variables. This gives a station specific per visit detection rate, and hence your likelihood of getting a false negative as a function of habitat type.

Assuming that we solve all these problems - we now have a biologically meaningful metric. We should have looked at all the assumptions that are inherent in this metric e.g. RBV, tested them, and made adjustments on the basis of biases.”

Following these recommendations, Dr. Gary White (Colorado State University) performed the necessary analyses, as shown in the appended report.

The objective of Dr. White’s analysis was essentially to determine whether some of the critical assumptions of the RBV approach were met, and to indicate any necessary adjustments. Of most importance was whether the probability of detecting occupancy behavior varies with habitat type. If, for instance, Murrelets were more easily detected in residual old-growth habitat than in unentered old-growth, this would cause a systematic bias in estimates of bird density (the numbers in unentered old-growth would be underestimated). If such a bias was found, and was found to be large, a correction factor would be calculable, and would be necessary before the RBV approach could be adopted.

Discussion of findings

Dr. White's analysis was successful in several important aspects. Notably he determined that the probability of detecting Murrelets did show substantive variation over time and space.

Of greatest importance was that p', the conditional probability of detecting occupancy on a given day, did show between habitat variance. As shown in White's Figure 2, p' is highest in old growth Redwood habitats, lower for residual forests, and lowest for habitat type O1/RD (mixed Redwood and Douglas Fir old-growth). However these differences themselves vary over the course of the season, as p' varies with Julian date.

Given that p' varies with habitat, RBV value might be overestimated in old-growth Redwood, and underestimated in residual and mixed old-growth forests. However this error rate is very small. The overall probability of detection of occupancy in a stand, x, is given by:

$$x = 1 - (1-p')^n$$

where n is the number of visits to a stand.

In Marbled Murrelet surveys, n is usually 8 (for simple detection of birds) or 20 (for determination of occupancy). From the above equation, we can estimate the maximum effect of variation in p' with habitat. For instance, in Figure 2, White shows that p' varies in mid season from values of .4 to .5 (approximately). From the above equation we can calculate that, when n = 8, in a habitat where p' is .4, 98.3% of occupied stands will be correctly detected. When p' is .5, the corresponding value is 99.6%. Hence the 'underestimation' of Murrelets when p' is 0.4 (relative to the number seen when p' is 0.5) = I- (98.3/99.6) or approximately 1.3%.

Such habitat effects, even at their most extreme value (mid-season) are thus shown to be of very small influence on estimates of occupancy, provided that n is on the scale indicated. At other times of the year, habitat effects will be even smaller (see Figure 2).

The most important result from White's analysis is thus that habitat specific effects on probability of detection of occupancy are small, provided that survey effort is adequate. Under these conditions, no systematic adjustment to RBV values appears justified, and the metric is thus supported in its existing form.

Also of interest from White's analysis is his calculation of a value of p' of 0.441. This value is substantively higher than that suggested for other parts of the range of the Marbled Murrelet (approximately 0.258). This suggests a higher degree of confidence in the accuracy of surveys carried out for this analysis than is generally the case.

<u>Application of results</u>

White's analysis suggests that correction factors for the RBV approach will produce only minor differences in estimation of Murrelet habitat. At this point, correction factors are thought to be unwarranted, and the assumptions of the RBV approach are generally supported.

However, this HCP has elected not to use the RBV approach. Of the different metrics available, the RBV approach is the least conservative, and has the maximum estimate for the numbers of birds protected on reserves. In keeping with the precautionary principle applied elsewhere in this HCP, the most conservative metric (number of occupied acres of habitat) has been used.

Logistic Regression Model of Detection Probability for Marbled Murrelets

Gary C. White
Department of Fishery and Wildlife Biology
Colorado State University
Fort Collins, CO 80523

4 May 1998

Executive Summary

A model of detection probability of marbled murrelets at standardized surveys of stations on Pacific Lumber Company and **nearby** timber lands is developed. Through the method of maximum likelihood, it is shown that the all the information on probability of detection is contained in the stations where murrelets were detected at least once during the year.

Logistic regression models of probability of detection of occupancy were developed. A simple model with constant probability of detection across stations and visits predicts the probability of detection per visit to be 0.441 (SE 0.019). This simple model is inadequate because of a general lack of fit, and because detection probabilities are shown to change across the survey season. Some of the variation in the data seems to be caused by the use of 81 different observers. However, the large number of observers precludes developing a parsimonious model based on observer effects.

The minimum AIC model was found to include Julian date (based on a 5th degree polynomial model), year (1992 through 1995), habitat type (3 types of timber canopy cover categories), and distance to the nearest body of water as measured on a 7.5 minute topographic quad. Inclusion of habitat type and distance to the nearest body of water is probably an artifact of the analysis in that each survey was classified as either occupied or not occupied. However, some stations were occupied by more than a single pair, causing increased probability of detection of occupancy. The inclusion of year may represent year-to-year differences in detection probabilities, possibly caused by weather differences, or year effects may represent differences in abundance of murrelets across years.

Unbiased estimates of the number of occupied stations can be obtained from the logistic regression model developed here with the Horvitz-Thompson estimator.

Two suggestions for further work are provided: develop a murrelet observability rating for each station included in the analysis, and conduct further work on estimating occupancy at stations with the general model developed here that includes both detection probability and probability of occupancy.

Introduction

In this report, a model of the detection probability of marbled murrelets ***(Brachyramphus marmoratus)*** is developed from standardized surveys conducted at survey stations on Pacific Lumber Company and nearby timber lands. Surveys were conducted with the protocols described by Ralph et al. (1994).

Statistical Model of Murrelet Detection

The statistical model assumed for this analysis is motivated from the work of Hunter et al. (1997) and Stauffer et al. (1995). Assume that the probability of a station being occupied is p, and the probability of detecting occupancy on survey i is p_i'. The p_i' are termed the detection probabilities. The p_i' are assumed independent in that the detection probability for a given survey is not affected by previous or subsequent surveys. This assumption is reasonable for marbled murrelets because surveys are passive, i.e., observers do not affect the behavior of the birds. To illustrate this model, the case with 3 surveys of each station will be presented. The status of a station must be one of the following 8 mutually exclusive possibilities:

Occupancy Detected on Surveys	Probability	Number of Stations
1, 2, 3	$pp_1'p_2'p_3'$	n_{111}
1,2	$pp_1'p_2'(1 - p_3')$	n_{110}
1,3	$pp_1'(1 - p_2')p_3'$	n_{101}
2,3	$p(1 - p_1')p_2'p_3'$	n_{011}
1	$pp_1'(1 - p_2')(1 - p_3')$	n_{100}
2	$p(1 - p_1')p_2'(1 - p_3')$	n_{010}
3	$p(1 - p_1')(1 - p_2')p_3'$	n_{001}
No detections	$p(1 - p_1')(1 - p_2')(1 - p_3') + (1 - p)$	n_{000}
Sum	1	n

where the subscript for the number of stations with each set of occupancy detections is 1 if occupancy is detected, and 0 otherwise. The number of stations surveyed is n, or the sum of the third column of the table. The sum of the 8 cell probabilities is 1, so that a multinomial distribution is defined. Note that the last cell, labeled “No detections,” consists of two terms: the probability that no detections are made at an occupied station, and the probability that a station is not occupied.

The likelihood function for the observed data is constructed as the product of the 8 cell probabilities, each raised to the power n_i. The maximum likelihood estimates of the 4 parameters can be solved numerically for a 2-survey problem. However, I have not been able to solve analytically the 3-survey problem, for the p_i', but have found that the estimator of $\hat{p}$ is

$$\hat{p} = \frac{n_{001} + n_{010} + n_{100} + n_{011} + n_{101} + n_{110} + n_{111}}{n(\hat{p}_1'(1 - \hat{p}_2')(1 - \hat{p}_3') + \hat{p}_2'(1 - \hat{p}_3') + \hat{p}_3')}.$$

For 2 surveys, the analytical estimates are:

$$\hat{p} = \frac{(n_{11} + n_{10})(n_{11} + n_{01})}{n\,n_{11}} = \frac{n_{11}}{n\,\hat{p}_1'\hat{p}_2'},$$

$$\hat{p}_1' = \frac{n_{11}}{n_{11} + n_{10}}, \text{ and}$$

$$\hat{p}_2' = \frac{n_{11}}{n_{11} + n_{01}}, \text{ with variances}$$

$$\hat{\text{var}}(\hat{p}_1') = \frac{\hat{p}_1'(1 - \hat{p}_1')}{n\,\hat{p}\,\hat{p}_2'},$$

$$\hat{\text{var}}(\hat{p}_2') = \frac{\hat{p}_2'(1 - \hat{p}_2')}{n\,\hat{p}\,\hat{p}_1'}, \text{ and}$$

$$\hat{\text{var}}(\hat{p}) = \frac{\hat{p}(1 - \hat{p})}{n} + \frac{\hat{p}}{n}\left[\frac{1 - \hat{p}_1'}{\hat{p}_1'}\right]\left[\frac{1 - \hat{p}_2'}{\hat{p}_2'}\right].$$

Note that the estimator for $\hat{p}$ reduces to a simple binomial estimator for $\hat{p}_1' = \hat{p}_2' = 1$, and its variance estimator similarly reduces to the standard binomial variance estimator.

These estimates demonstrate that all the information to estimate the detection probabilities is provided only by stations where occupancy was detected at least once, and that stations where occupancy was never detected provide no information on the detection probabilities. Thus, the only data needed to estimate detection probabilities is provided by the stations where at least one survey resulted in occupancy being detected at the station. The detection model developed from only stations with occupancy is fully efficient, with no loss of information by discarding stations where murrelets were never detected.

Statistical Methods

Logistic regression using PROC GENMOD from SAS Institute Inc. (1997) was used to develop a model to estimate detection probabilities from data on stations with occupancy at least once during the year. The dependent variable is whether or not a station was determined to be occupied during a visit. The independent variables described in Table 1 were considered in developing a model of detection probability. The linear model used in a logistic regression is defined as

$$\text{logit}(p') = \log_e\left(\frac{p'}{1 - p'}\right) = \beta_0 + \beta_i \text{ (covariate } i),$$

with the model extendable to multiple covariates. The method of maximum likelihood is used to estimate the unknown parameters $(\beta_0, \beta_1, \ldots)$ with the binomial distribution assumed as the error structure on the residuals in the model. This model can be considered as a special case of the statistical model described above, where the likelihood is partitioned into 2 parts. Only the portion containing the detection probabilities is analyzed with logistic regression. Multiple surveys for a station result in a detection probability for each survey. By including the date of the survey as part of the model predicting detection probability, the temporal nature of the detection probabilities across surveys is maintained.

Akaike's information criterion (AIC) (1973) was used to determine the most parsimonious model fitting the data. The "best" model is the model with the smallest AIC value. AIC is defined as $-2 \times \log_e(\text{Likelihood}) + 2 \times$ number of parameters. The $\log_e((\text{Likelihood})$ portion represents the goodness of fit of the model to the data. The better the fit, the smaller the term $-2 \times \log_e((\text{Likelihood})$. However, the second term represents a penalty for the number of parameters. The more parameters included in the model, the better the fit, but the bigger the penalty.

Goodness-of-fit of the logistic regression model is assessed with the deviance [deviance $= -2 \times \log_e((\text{Likelihood})$] divided by the degrees of freedom. The expected value of this criterion is unity. Models with deviance/df > 1 suggest over-dispersion (i.e., extra binomial variance) exists. That is, the model does not fully explain the observed variation in the detection probabilities. Heterogeneity of detection probabilities still exists, requiring additional explanatory variables to account for the observed variation of the detection probabilities.

The first cut at model selection was to include each variable in Table 1 alone in a logistic regression model. Based on these results, more complex models including multiple covariates were developed. Only models that seemed biologically reasonable a priori were considered.

Table 1. Independent variables used to develop a detection model with logistic regression for marbled murrelets. The column labeled AIC is the AIC value for a model with just this variable included.

Variable Name	Description	AIC
JDATE	Julian day, i.e., day of the year, or number of days since 1 January	95 1.2026
YEAR	Year surveys were conducted (1992, 1993, 1994, 1995)	958.0297
DIST2CST	Distance (m) to coast	966.5979
DIST2H20	Distance (m) to nearest water on a 7.5 min topographic quad	962.9285
DIST20CC	Distance (m) to nearest station with occupancy behavior	966.9158
DIST20G	Distance (m) to nearest old-growth patch	968.6828
DIST2RES	Distance (m) to nearest residual patch	964.5022
ELEV	Elevation (m)	967.9372
FRACT	Fractal-shape index of associated patch	968.1091
HABTYPE	Habitat type of station or within 100 m (12 types)	973.4910
MAJ	Nearest timber class within 100 m of station	988.5136
AREA	Area of habitat patch at station or within 100 m	968.5648
PERIM	Perimeter of habitat patch at station or within 100 m	968.7017
SHAPE	Shape index of habitat patch at station or within 100 m	967.3723
SLOPE	Slope (%) at station	965.8509
TRUMAJ	Timber or vegetation classification at station	990.2350
CANPCODE	Canopy code	968.7346

Results

With no independent variables included in the model, the probability of detection is estimated as 0.441 (SE 0.019) with an AIC value of 966.7426. However, this model shows considerable over-dispersion, i.e., extra binomial variation. The ratio of the deviance to the degrees of freedom for this simple model is 964.7426/702 = 1.3743, strongly suggesting that the probability of detection is not constant across stations or visits. That is, the assumption that the detection probability remains constant across stations and surveys is invalid.

One explanation for the observed over-dispersion is variation across sites because of variation in murrelet density. However, when a variable consisting of the maximum number of mm-relets detected as occupying a station (MAXOCC) was used as a predictor variable, the deviance/df ratio declined only a small amount to 1.3677. This variable is a significant predictor of the probability of detecting murrelets at a stations ($P < 0.001$), but **is** only measurable after-the-fact, so is not useful in developing a detection function. That **is,** the use of this variable in a detection function is circular, and is only used here to evaluate its impact on deviance. The small decline in deviance/df suggests that the cause of the over-dispersion is not due to variation in murrelet density.

Most likely, part of this over-dispersion is explained by observers. Eighty-one different observers were reported in the data base, with significant differences between observers ($x^2 = 104.925$, df = 80, $P = 0.032$). However, a parsimonious model to remove this over-dispersion from observers is not possible because of the large number of observers, and because I know of no way to objectively pool observers into categories. Thus, some amount of over-dispersion will remain in the data even when the covariates from Table 1 are used to model detection because of the observer differences.

AIC results for each variable in Table 1 are shown in the table. The variables JDATE, YEAR, DIST2H20, and DIST2RES were considered further **in** developing a model of detection probability because these variables have the smallest AIC values. In particular, JDATE provided the smallest AIC value of all the single-variable models considered. Because detection probabilities were thought to change over the period when surveys are conducted, I evaluated polynomial models of JDATE.

Polynomial Power	Number of Parameters	AIC
1	2	95 1.2026
2	3	947.5 148
3	4	946.9036
4	5	941.8112
5	6	939.9904
6	7	94 1.9904

These results suggest that a 5^{th} degree polynomial model is required to model the effects of Julian date on detection probability. A plot of the resulting function is shown in Figure 1, with 95% confidence intervals on the predicted value included. The maximum detection probability is on day 202, or July 21. Actual surveys were conducted from Julian day 104 to 216, so the graph encompasses the survey dates.

The habitat variables TRUMAJ and MAJ are ranked poorly by AIC because of the large number of parameters required to model them (27 and 24, respectively). Likewise, HABTYPE requires 9 parameters in the model. However, inspection of the frequency of each habitat type in the data where at least one occupancy was detected during the year suggests that the habitat types should be further restricted because so few observations occurred for 5 of the habitat types:

HABTYPE Code	Habitat Patch Description	Frequency	Percent
>100m	Habitat patch > 100 m from station	16	2.3
O1/R	Old growth with >50% cover of redwoods	121	17.2
Ol/RD	Old growth with >50% cover of redwoods and Douglas fir	265	37.7
01/D	Old growth with >50% cover of Douglas fir	4	0.6
02/RD	Old growth with <50% cover of redwoods and Douglas fir	8	1.1
R1/R	Residual growth with >50% cover of redwoods	8	1.1
R2/R	Residual growth with <50% cover of redwoods	258	36.7
R2/RD	Residual growth with <50% cover of redwoods and Douglas fir	14	2.0
R2/D	Residual growth with <50% cover of Douglas fir	9	1.3

Therefore, only habitat type 01/R, Ol/RD, and R2/R were used in further analyses. With this reduced data set, habitat type becomes an important variable (Table 2) because only 2 parameters are required to model the habitat type variable, as opposed to 8 parameters for the full list of habitat types.

From Table 2, the most parsimonious model is determined to be Julian date as a 5^{th} degree polynomial, year, habitat type, and distance to nearest water body. The numerical estimates and other output from SAS are provided in Appendix A.

In Figure 2, this detection model is plotted for the 3 habitat types and Julian date, assuming that year is 1992 and distance to nearest water body is 0. Distance to nearest water caused a slight decline in detection probability as distance to water increased. The probability of detection is greatest for habitat type 01/R, lowest for habitat type O1/RD, and intermediate for habitat type R2R. In Figure 3, this same detection model is plotted for the 4 years and Julian date, assuming that habitat type is 01/R and distance to nearest water body is 0.

The minimum AIC model is still not a suitable model for estimating detection probabilities because the deviance/df value for this model is 825.8577/632 = 1.3067. This high value suggests that the information to estimate detection probabilities **is** not contained in the list of variables in Table 1, and that additional variables are needed to estimate station and visit-specific detection probabilities.

Table 2. Multiple covariate models used to develop a detection model with logistic regression for marbled mm-relets.

Model	Number of Parameters	AIC
Intercept Only	1	885.7837
JDATE^5 (5th degree polynomial of Julian date)	6	865.6452
HABTYPE	3	883.1654
JDATE^5 + YEAR	9	854.8195
JDATE^S + YEAR + HABTYPE	11	850.2556
JDATE^S + YEAR + DIST2H20	10	853.1801
JDATE^5 + YEAR + DIST2RES	10	853.1801
JDATE^S + YEAR + HABTYPE + DIST2H20	12	849.8577
JDATE^5 + YEAR + HABTYPE + DIST2RES	12	850.9000
JDATE^5 + YEAR + HABTYPE + DIST2H20 + DIST2RES	13	851.1550
JDATE^5 + YEAR + DIST2H20 + DIST2RES	11	852.7899
JDATE^S + HABTYPE	8	859.5311
JDATE^5 + HABTYPE + DIST2H20	9	858.1918
JDATE^S + HABTYPE + DIST2RES	9	859.1944
JDATE^5 + HABTYPE + DIST2H20 + DIST2RES	10	858.9042

Discussion

The inclusion of habitat type and distance to nearest water body in the minimum AIC model suggests that these variables influence detection probabilities. However, I am inclined to think that these variables are more likely to influence the probability that mm-relets occupy a station. The reason that these variables get included in the model is that occupancy of a station is treated as either yes or no, but in reality, occupancy occurs in degrees. Some stations have more than 1 pair of murrelets in occupancy, which results in increased probability of detection. As a result, the model predicts greater detection probabilities, when in reality, habitat effects are increasing detection probability only because of a greater number of murrelets at the station. Credibility for this argument is provided when the variable MAXOCC, consisting of the maximum number of birds detected occupying the station during the year, is included in the model. The variable DIST2H20 is no longer significant (P = 0.476), suggesting that the influence it provided in the best detection model is replaced by the MAXOCC variable. MAXOCC is clearly correlated with the suitability of the station for murrelets.

Inclusion of year in the minimum AIC model suggests year-to-year differences in detection probabilities. This difference may be due to year-to-year differences in occupancy rates reflecting murrelet abundance, and be related to the problem discussed in the preceding paragraph concerning habitat variables. However, the effect of year may also **be** related to differences in weather from year to year (e.g., fog, drizzle, or other weather that causes observability problems), causing differences in detectability. Regardless of which of the above reasons explains why year was included in the model, the fact that year is an important predictor of the detection probability implies that examining trends in occupancy rate requires differential corrections by year, particularly given the range (>0.2) of variation in detection probabilities.

An alternative explanation for why year is included in the model is that year-to-year differences in the areas surveyed could explain the effect, That is, only high quality mm-relet habitat might have been surveyed in 1995, compared to poorer quality habitat during 1992-1994, causing the apparent year-to-year observed differences.

Plots of the detection probability as a function of Julian date consistently show a strong decline in detection probability at the beginning of the survey season. This decline may suggest that surveys should be started earlier in the year. However, a confounding of only surveying high occupancy stations early in the season could also cause this initial early decline. More interpretation of this early decline in detection probability is warranted.

The high deviance/df value for the minimum AIC model suggests that additional information is required in the model to estimate detection probabilities. I think that a more appropriate variable could be constructed through observer savvy. That is, observers have some sense of how detectable murrelets are at a station, given that murrelets are present. A team of observers could independently rate a series of stations from low (1) to high (10) detection, where the stations are selected to maximize the range of observability. The mean of the team's ratings would be assigned to the station, and future observers could then calibrate their observability scale by visiting the stations. In this way, observers could be trained to estimate observability reliably and consistently, and an observability rating could be assigned to each station in future surveys. I think that this observability variable, although somewhat subjective, would result in a better model of detection when combined with JDATE than the model developed here.

The logistic regression developed in this paper can be used with the approach of Steinhorst and Samuel (1989) and Samuel et al. (1987) to estimate the total number of occupied stations, even when stations are only surveyed one time. A modified Horvitz-Thompson estimator (Horvitz and Thompson 1952, Cochran 1977:259-261) can be used. Suppose that n stations are surveyed, with the results for each station coded as $y_i = 1$ for station *i* when a detection is made, and $y_i = 0$ when no detection is made. Then, the Hot-&-Thompson estimator of the number of occupied stations is $\hat{Y}_{HT} = \sum_{i=1}^{n} \frac{y_i}{\hat{p}_i}$, where $\hat{p}_i$ is the probability of a detection at station ***i*** as predicted by the logistic regression equation. Full details and the variance estimator are provided by Steinhorst and Samuel (1989). Note that the variance estimator given in Cochran (1977:260) is not correct for this situation because Cochran assumes that the detection probabilities (p_i) are known exactly, and are not estimates as is the case here. Steinhorst and Samuel (1989) provide for the additional variation from the estimated detection probabilities in their formulae.

The detection models proposed by Max et al. (Undated) are essentially special cases of the model developed here. Their models assume that detection probabilities are constant across stations and date, and hence correspond to the case of a logistic regression with only an intercept. Further, the estimators they propose are essentially equivalent to the maximum likelihood estimators developed by Darroch (1958) (also model $M_0$of Otis et al. (1978:21-24)) for the murrelet example, where stations are visited even after detections are made at the station, and Zippin (1956, 1958) (model M_b of Otis et al. (1978:28-32)) for the spotted owl example where visitation stops once a detection is made. Both of these estimators were originally developed to estimate population size *(N),* which in the situation considered here corresponds to the number of occupied stations. Max et al. (Undated) consider in some detail the degenerate case where the estimated number of stations occupied exceeds the number of stations surveyed.

Suggestions for Further Research

1. Develop and implement an observability rating for each murrelet survey station, and incorporate this rating into a detection model.

2. Further work on estimating the presence of murrelets at survey stations should incorporate the detection model developed here (or an improved model as suggested above) as part of the general statistical model that includes the probability of a station being occupied. This research would require software to implement numerically the general model of murrelet detection that includes both the probability of murrelet occupancy and probability of detection.

Suggestions for Improving Surveys

1. Decrease the variation across observers by reducing the number of observers, and having each of them do more surveys.

2. Incorporate a variable that directly relates to murrelet observability at each station.

3. Further investigate the reason for the initial decline of the detection probability, and if this decline is real, initiate surveys earlier in the year.

Acknowledgments

S. L. Miller and C. J. Ralph provided access to the data to develop this model. S. L. Miller was a particularly gracious host during my visit to Redwood Sciences Laboratory to discuss the data set. K. P. Bumham provided advice on specific aspects of maximum likelihood estimation. A. B. Franklin provided some critical references and discussion on the techniques, and made some suggestions for improvement of this document.

Literature Cited

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. Pages 267-281 in B. N. Petrov, and F. Csaki, (Eds.) Second International Symposium on Information Theory. Akademiai Kiado, Budapest.

Cochran, W. G. 1977. Sampling Techniques. 3rd. Ed. John Wiley & Sons, New York, N.Y. 428 pp.

Darroch, J. N. 1958. The multiple-recapture census: I. Estimation of a closed population. Biometrika 45(3/4):343-359.

Hunter, J. E., K. N. Schmidt, H. B. Stauffer, S. L. Miller, C. J. Ralph, and L. Roberts. 1997. Status of the Marbled Mm-relet in the inner north coast ranges of California: final results. Unpublished report. 42 pp.

Max, T. A., J. A. Baldwin, and H. T. Schreuder. Undated. Sampling threatened and endangered birds in the Pacific Northwest. USDA Forest Service Unpublished report. 21 pp.

Ralph, C. J., S. K. Nelson, M. M. Shaughnessy, S. L. Miller, and T. E. Hamer. 1994. Methods for surveying for marbled murrelets in forests: a protocol for land management and research. Pacific Seabird Group, Redwood Sciences Laboratory, Arcata, CA.

Samuel, M. D., E. 0. Garton, M. W. Schlegel, and R. G. Carson. Visibility bias during aerial surveys of elk in northcentral Idaho. Journal of Wildlife Management 51:622-630.

SAS Institute Inc. 1997. SAYSTAT® software: changes and enhancements through release 6.12. SAS Institute Inc., Gary, NC, USA. 1167 pp.

Stauffer, H. B., K. N. Schmidt, J. E. Hunter, S. L. Miller, C. J. Ralph, and T. M. Jimerson. 1995. The range and distribution of the marbled murrelet ***(Brachyamphus marmoratus)*** at far inland sites in the northern California: a sampling design. Unpublished report. Not consecutively paginated.

Steinhorst, R. K., and M. D. Samuel. 1989. Sightability adjustment methods for aerial surveys of wildlife populations. Biometrics 45:415-425.

Zippin, C. 1956. An evaluation of the removal method of estimating animal populations. Biometrics 12: 163-l 89.

Zippin, C. 1958. The removal method of population estimation. Journal of Wildlife Management 22:82-90.

Figure 1. Marbled murrelet detection **function** with 95% confidence intervals on the predicted value ("mean") modeled as a 5" degree polynomial of Julian date.

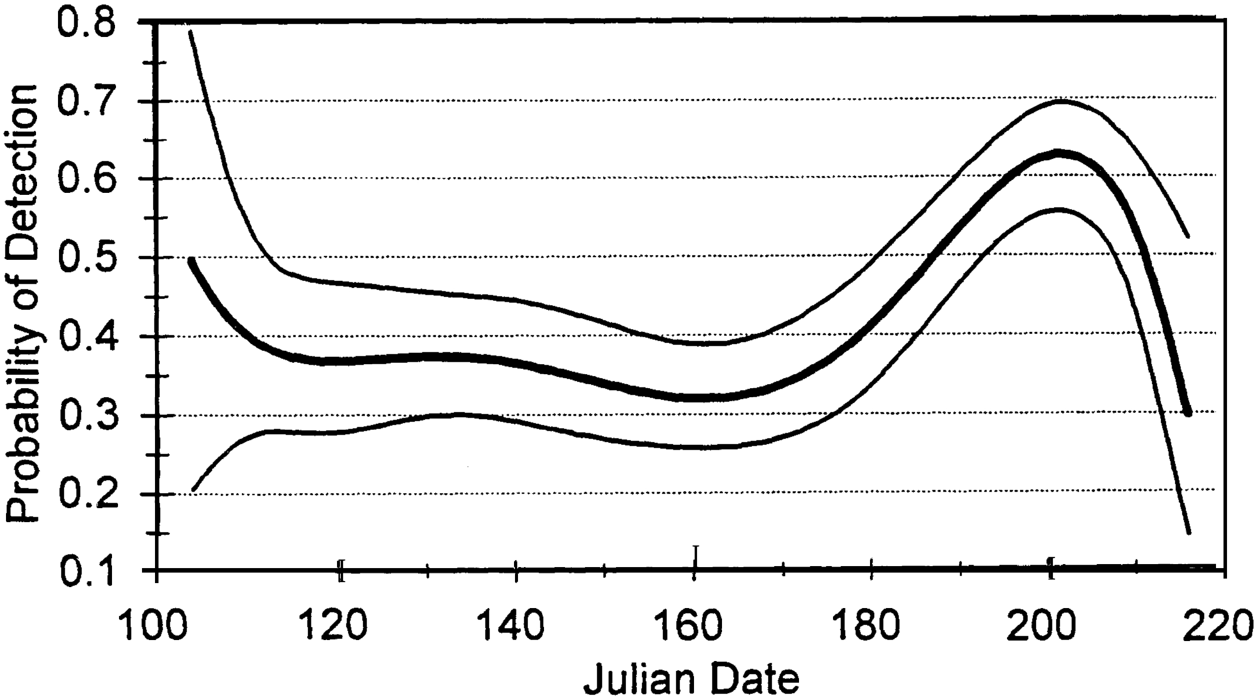


Figure 2. Predicted marbled murrelet detection probability from the minimum AIC model. Julian date is modeled as a 5^{th} degree polynomial for the 3 habitat types that represent 95% of the data set. Year is set to 1992, and distance from nearest water is set to zero.

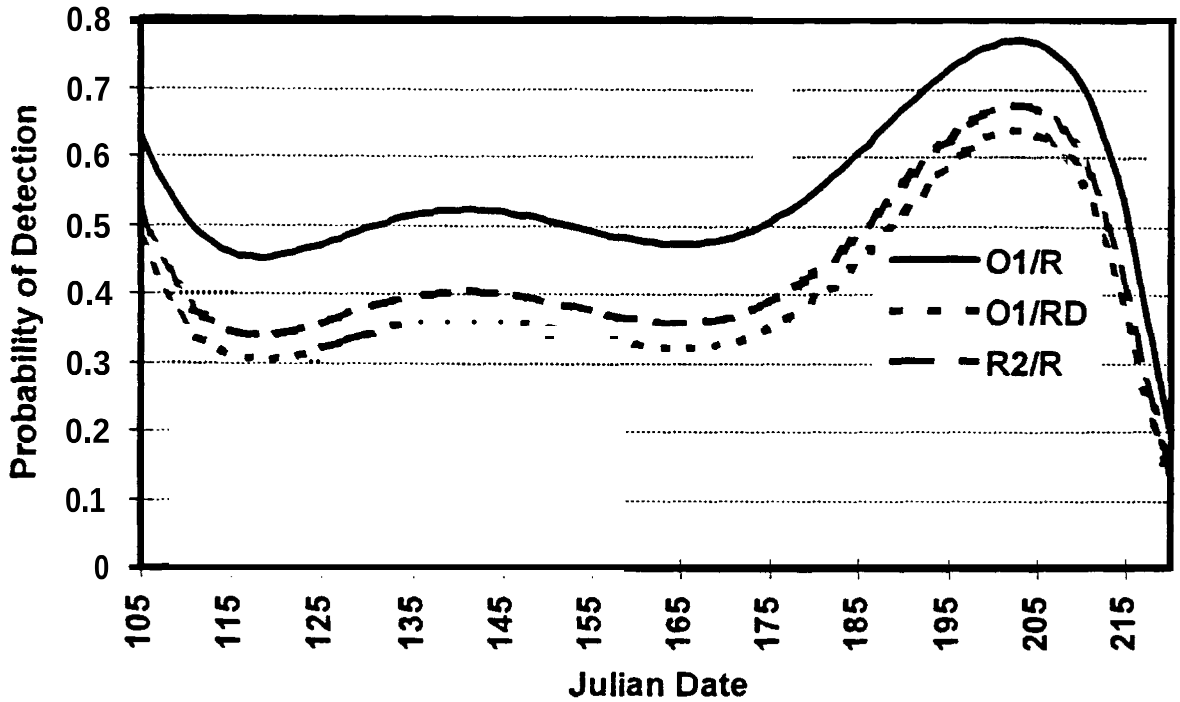
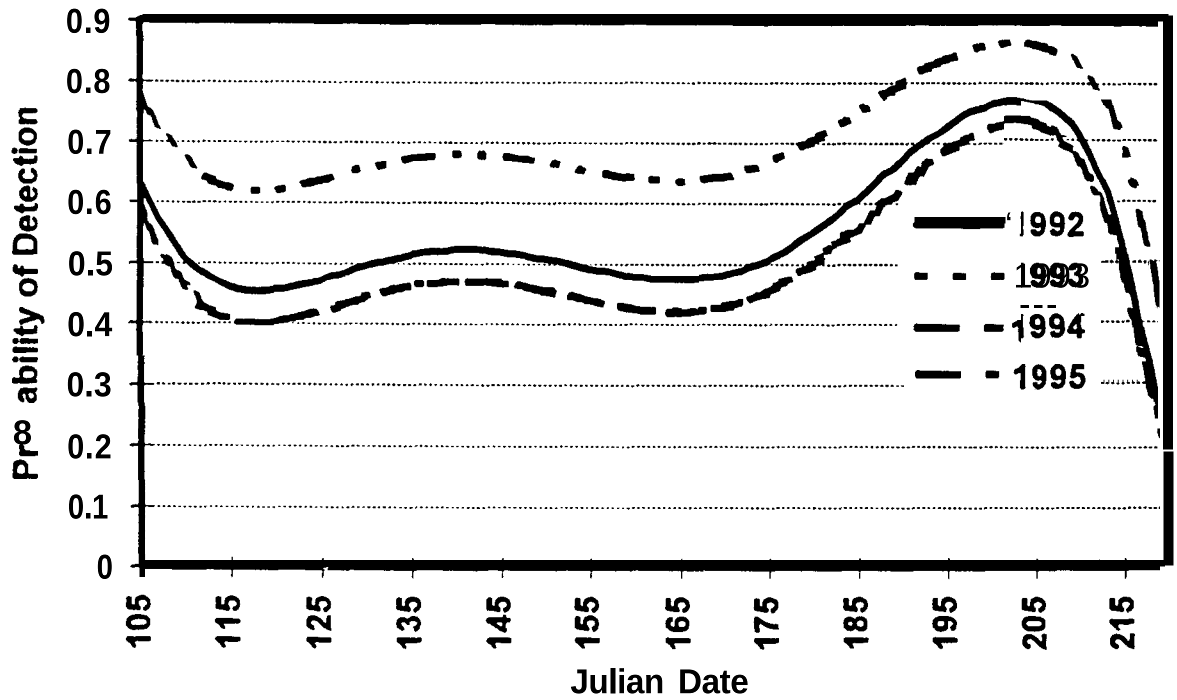


Figure 3. Predicted marbled murrelet detection probability from the minimum AIC model. Julian date is modeled as a 5^{th} degree polynomial for the 3 habitat types that represent 95% of the data set. Habitat type is set to O1/R, and distance from nearest water is set to zero. The detection probabilities for 1993 and 1994 are nearly identical, so that the lines overlay.



Appendix A - SAS Output for minimum AIC Model

Detection model for Year, Habtype, Distance to H20, and Julian Date

The GENMOD Procedure

Model Information

Description	Value	Label
Data Set	WORK.DETECT	
Distribution	BINOMIAL	
Link Function	LOGIT	
Dependent Variable	VISITDET	Occupancy Detection
Dependent Variable	ONE	
Observations Used	644	
Number Of Events	284	
Number Of Trials	644	

Class Level Information

Class	Levels	Values
HABTYPE	3	O1/R O1/RD R2/R
YEAR	4	1992 1993 1994 1995

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	632	825.8577	1.3067
Scaled Deviance	632	825.8577	1.3067
Pearson Chi-Square	632	645.3980	1.0212
Scaled Pearson X2	632	645.3980	1.0212
Log Likelihood	.	-412.9289	

Analysis Of Parameter Estimates

Parameter		DF	Estimate	Std Err	ChiSqua·	Pr>Chi
INTERCEPT		1	846.5581	409.6540	4.2705	0.0388
YEAR	1992	1	-0.6721	0.3016	4.9651	0.0259
YEAR	1993	1	-0.8708	0.2701	10.3940	0.0013
YEAR	1994	1	-0.8752	0.2501	12.2443	0.0005
YEAR	1995	0	0.0000	0.0000	.	.
HABTYPE	01 /R	1	0.4716	0.2389	3.8979	0.0483
HABTYPE	01 /RD	1	-0.1623	0.1900	0.7296	0.3930
HABTYPE	R2 /R	0	0.0000	0.0000	.	.
DIST2H2O		1	-0.0009	0.0006	2.3542	0.1249
JDATE		1	-28.4324	13.3079	4.5647	0.0326
JDATEP		1	377.2093	170.6939	4.8835	0.0271
JDATE3		1	-247.0112	108.0863	5.2227	0.0223
JDATE4		1	79.8244	33.8028	5.5786	0.0182
JDATE5		1	-10.1815	4.1791	5.9358	0.0148
SCALE		0	1.0000	0.0000	.	

NOTE: The scale parameter was held fixed.

LR Statistics For Type 3 Analysis

Source	DF	ChiSquare	Pr>Chi
YEAR	3	14.3340	0.0025
HABTYPE	2	6.8960	0.0318
DIST2H2O	1	2.3979	0.1215
JDATE	1	4.5591	0.0327
JDATEP	1	4.8866	0.0271
JDATE3	1	5.2357	0.0221
JDATE4	1	5.6012	0.0179
JDATE5	1	5.9736	0.0145